



IN THE

UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS: Olorenshaw et al.

SERIAL NO.: 09/369,031

FILED: August 4, 1999

TITLE: System And Method For Speech Recognition
Using An Enhanced Phone Set

EXAMINER: Knepper, D.

ART UNIT: 2654

ATTY DKT NO: 3188.01/1144

RECEIVED

MAR 23 2004

Technology Center 2600

CERTIFICATE OF MAILING

I hereby certify that this paper is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450, on the date printed below:

Date: 3/16/04



Gregory J. Koerner

Amendment And Response

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Dear Sir:

In response to the Non-Final Office Action mailed December 17, 2003,
please reconsider the above-identified Application in light of the following remarks
and amendments.

In the Specification:

Please replace the paragraph found on page 8, lines 14 through 24, with the following substitute paragraph:

HMM 212 utilizes the feature vectors and endpoint information to recognize speech. HMM 212 analyzes the feature vectors between endpoints to produce a series of phones, which are input to phonetic dictionary 214 via path 328. Phones, also known as phonemes or phone-like units (plu), represent basic sound units of a spoken language. HMM 212 is a software module executed by CPU 128 to match feature vectors from a portion of speech data to one of the phones. ~~Hidden Markov Models are further discussed in Automatic Speech Recognition: The Development of the SPHINX System, by Kai Fu Lee, Kluwer Academic Publishers, Boston, 1989, which is hereby incorporated by reference.~~

Please replace the paragraph found on page 10, lines 4 through 14, with the following substitute paragraph:

The transformed phone dataset is output to HMM training process 418 via path 428. HMM training process 418 preferably utilizes the transformed phone dataset to produce a HMM 212 for a particular speech recognition application. HMM training process 418 typically creates a software model for each phone in the transformed phone dataset, and then each model is taught which speech data correspond to the model's phone using statistical learning techniques. ~~Hidden Markov Model training is further discussed in "Speaker Independent Phone Recognition Using Hidden Markov Models," by Kai Fu Lee and Hsiao Wuen Hon, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 37, No. 11, 1989, pp. 1641-1648, which is hereby incorporated by reference.~~

Please replace the paragraph beginning on page 12, line 26, and ending on page 13, line 5, with the following substitute paragraph:

Transcription 714 captures many details of the acoustic-phonetic content of speech token 712. Composite-phone /bcl<h/ represents a b-closure with devoicing at the end of the phone. Composite-phone /s=z/ represents an s-fricative with voiced frication (hissing sound) affecting the entire phone. Composite-phone /kcl=v/ represents a k-closure that is completely voiced, and composite-phone /k=v/ represents a k-stop that is completely voiced. Composite-phone /v>tcl/ represents a t-closure that is voiced at the beginning of the base-phone. Composite-phone /t<a=h/ represents a t-stop with an epenthetic vowel process affecting the end of the base-phone with complete devoicing. ~~A description of these and other phonetic terms may be found in A Course in Phonetics, by Peter Ladefoged, Harcourt Press, 1982, which is hereby incorporated by reference.~~

Please replace the paragraph found on page 14, lines 5 through 17, with the following substitute paragraph:

Referring now to FIG. 8(a), a table for one embodiment of an extended base-phone set 810 is shown, according to the present invention. The FIG. 8(a) embodiment of extended base-phone set 810 of the present invention includes a subset 812 and a subset 814. Subset 812 includes the commonly-used TIMIT base-phone set. The TIMIT base-phone set includes base-phones that represent stops and affricates, fricatives, nasals, semivowels and glides, diphthongs and syllabics, vowels, silence, closure, pause, and stress. ~~The TIMIT base-phone set is further described in "Speech Database Development: Design and Analysis of the Acoustic-Phonetic Corpus," by L.F. Lamel et al., Proc. DARPA Speech Recognition Workshop, February 1986, pp. 100-109, which is hereby incorporated by reference.~~ The TIMIT speech corpus database is available on CD-ROM from the National Technical Information Service (www.ntis.gov).